An evaluation of the caught being good game with an adolescent student population

Journal of Positive Behavior Interventions

An Evaluation of the Caught Being Good Game with an Adolescent Student Population

Journal:	*Journal of Positive Behavior Interventions*
Manuscript ID	JPBI-19-112.R3
Manuscript Type:	Empirical Research
Keywords:	adolescent < Age group, classroom-based < Studies, single-case designs < data analysis < Studies, positive behavior < Support(s)

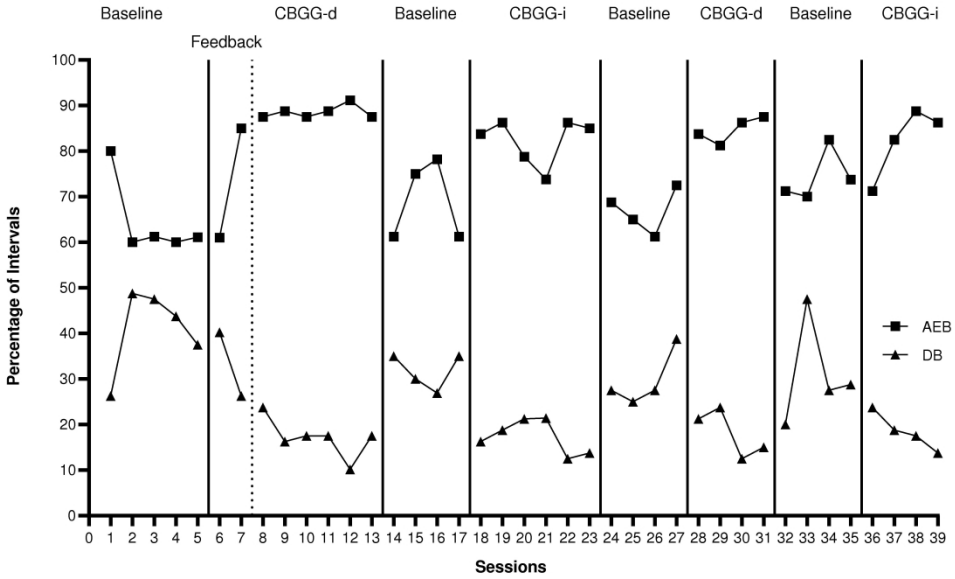

Figure 1. Percentage of intervals with AEB and DB across study phases

214x137mm (600 x 600 DPI)

Abstract

The current study investigated the Caught Being Good Game (CBGG), for use with an adolescent student population. The CBGG is a positive variation of the Good Behavior Game (GBG), a popular group contingency intervention in classroom management literature. In this positive version, teams of students receive points for engaging in desirable behavior, rather than marks for breaking class rules. Research on the CBGG has garnered empirical interest in recent years however there is little published research on the game with adolescent populations. The current study investigated if visual feedback displayed on a scoreboard during the CBGG is a necessary part of the game. This was examined by implementing the game both with and without overt visual feedback, using an ABACABAC reversal design. Academically engaged behavior and disruptive behavior were monitored. The CBGG was effective in both formats, leading to increases in academically engaged behavior and decreases in disruptive behavior in the participating class group. This suggests that perhaps immediate visual feedback is not an essential component of the CBGG for adolescent, mainstream students. This may be a time-saving measure for teachers wishing to implement the game. Students and their teacher rated the game favorably on social validity measures.

An Evaluation of the Caught Being Good Game with an Adolescent Student Population

Interdependent group contingency interventions offer an evidence-based, class-wide solution to challenging behavior in the classroom (Maggin et al., 2012; Stage & Quiroz, 1997). An interdependent group contingency is when a reward is given based on every member of a group or team meeting a certain performance criterion (Litow & Pumroy, 1975). The Good Behavior Game (GBG; Barrish et al., 1969) is a popular game-based interdependent group contingency applied in the classroom management literature (Bowman-Perrott et al., 2016; Flower et al., 2014; Tingstrom et al., 2006). The GBG involves dividing a class group into teams and providing marks throughout the game to teams on which a

member breaks a class rule. The aim is for the team to remain under a certain criterion of marks by the end of the game to receive a prize. The GBG has been altered recently to align with more positive classroom management practices. Rather than the provision of marks, teachers award points to teams who follow class rules. Teams meeting or surpassing a specified criterion of points at the end of the game are eligible for a prize. This positive variation is often called the 'Caught Being Good Game' (CBGG; Wahl et al., 2016; Wright & McCurdy, 2012), although some studies have called it the GBG (e.g., Groves & Austin, 2017) or the GBG-reinforcement (Tanol et al., 2010). In the current paper, the CBGG will be used to refer to all instances of a positive version of the GBG (involving the provision of positive points for rule-following), and the GBG will refer to the traditional version (involving the provision of negative marks for rule-breaking). The GBG and CBGG have been successful in reducing challenging and disruptive classroom behavior (Bowman-Perrott et al., 2016). Both versions have been effective in kindergarten (Donaldson et al., 2015; Tanol et al., 2010) and elementary school settings (Lannie & McCurdy, 2007; Nolan et al., 2014; Wahl et al., 2016; Wright & McCurdy, 2012) and the GBG has also been used with older (adolescent) students (Kleinman & Saigh, 2011; Mitchell et al., 2015).

Early evaluations of the CBGG yielded positive results. For example, Robertshaw and Hiebert (1973) demonstrated the efficacy of the CBGG in a first-grade classroom, awarding teams points in the form of tokens for good behavior. Improvements were noted in one particularly disruptive student's behavior and activity completion for the whole class increased. In other early research, implementation of the CBGG led to decreases in disruptive library behavior in a fourth-grade class (Fishbein and Wasik, 1981) and increases in appropriate social behavior across three physical education (PE) classes (Patrick et al., 1998). Patrick et al. (1998) also included a response-cost component whereby teams could lose

points for inappropriate behavior. Despite a paucity of research on the CBGG throughout the late 80s, 90s and early 2000s, there has been a resurgence in research in recent years.

In recent times, the CBGG has successfully been used to target a variety of behaviors; the reduction of rule violations (Tanol et al., 2010), increase of on-task behavior (Pennington & McComas, 2017) and the combined decrease in disruptive behavior and increase in academically engaged/on-task behavior (Wahl et al., 2016; Wright & McCurdy, 2012). It has been effective with mainstream classes using technology in its implementation (Ford, 2017; Lynne et al., 2017) and with students with emotional and behavioral disorders (Groves & Austin, 2017). Tanol et al. (2010) compared the CBGG to the GBG with response-cost in a kindergarten class group. In the CBGG condition, student teams gained stars for rule-following, whereas in the GBG response-cost condition, each team began the game with four stars, and stars were removed for rule-breaking. Teams with three stars at the end of the game were eligible for a prize. The participating teachers preferred the CBGG to the GBG response-cost. Other studies have compared the CBGG with the traditional GBG. Wright and McCurdy (2012) observed decreases in disruptive behavior and increases in on-task behavior across both game versions in two classrooms (kindergarten and fourth-grade). Wahl et al. (2016) followed up on this study, similarly demonstrating that the CBGG was just as effective as the GBG in targeting engagement and disruption across four classroom settings (kindergarten x2, first/second-grade mixed and second-grade). Notably, the researchers had the teacher record points discretely throughout the class and announce them to students at the end when the game was finished. Despite promising results, neither intervention was withdrawn with a return to baseline phase, limiting the conclusions which can be drawn.

Few studies have investigated the CBGG with adolescent students. One exception by Conklin et al. (2017) evaluated a group contingency game similar to the CBGG with seventh-grade students as part of the larger Class-wide Function-related Intervention Teams (CW-

FIT) program. The intervention was effective in targeting on-task behavior and compliance across two seventh-grade groups. Isolating the CBGG for examination with adolescent students is therefore an important line of inquiry for future research. In a doctoral thesis, Ford (2017), showed that the CBGG, when implemented in conjunction with ClassDojo technology, led to a decrease in disruptive behavior and increase in academic engagement across four seventh and eighth-grade classes. ClassDojo is an interactive platform which allows teachers to provide and display individual student or team points on an interactive whiteboard throughout the school day ("ClassDojo", 2019). Of the studies discussed, only one evaluated the CBGG alone with students older than fourth-grade (Ford, 2017), despite evidence that the traditional GBG has been implemented successfully with students at more advanced grade levels (Kleinman & Saigh, 2011; Mitchell et al., 2015). There is therefore a gap in our knowledge on the efficacy of the CBGG with adolescents and one of the main aims of the current research was to address this issue. It is also evident that there have been differences in methods of feedback used during the CBGG. Some studies have elected for teachers to record points discretely (e.g., Wahl et al., 2016) and some have had teachers record points in real time, on a public display (e.g., Lynne et al., 2017). In other studies, this element is ambiguous, with authors not explicitly stating how and when points are recorded publicly (Wright & McCurdy, 2012). A second aim of the current study was therefore to examine the method of feedback delivery as a variable component within the CBGG.

There have been a number of component analyses of the GBG which have evaluated the use of feedback during the game, however the work has not been conclusive and there is still much more to learn about this variable component. Medland and Stachnik (1972) evaluated the necessity of feedback during the GBG by putting it in place with two groups of fifth-grade students. Rather than feedback taking the form of marks on the board, feedback was delivered via a light system operated by observers and visible to teams of students. This

light was changed from green to red for 30 s contingent on a member of the team breaking a rule. After applying the full GBG-package with the class, the researchers withdrew the game. They then assessed whether a phase with rules only and a phase with rules and lights only could be effective in maintaining reductions in disruptive behavior. Harris and Sherman (1973) similarly conducted a component analysis of the GBG and attempted to delineate whether the public posting of marks on the board during the GBG was an essential component for the game's efficacy. The authors concluded that it was not essential as the level of disruptive behavior was low when the game was in place with no feedback. A flaw in both Harris and Sherman (1973) and Medland and Stachnik's (1972) studies is the presentation of the full GBG-package before manipulation of feedback elements. There is a strong possibility that the GBG-package phase led to crossover effects in the ensuing phases.

Foley et al. (2019) addressed this limitation by evaluating GBG components in a preschool setting before and after implementation of the whole GBG-package. Rules, a criterion of marks, feedback (in the form of publicly displayed marks) and reinforcement (both contingent and non-contingent) were examined as separate components in order to identify crucial components of the game. Each component was added one by one across several phases, culminating in the application of the whole GBG-package. The authors found that the GBG-package was necessary to produce meaningful decreases in disruptive behavior. After exposure to the GBG-package, rules, feedback, a criterion and non-contingent reinforcement demonstrated similar effects in keeping disruptive behavior low. Although the authors examined the feedback component here, they did not examine whether the game was successful without feedback but with a criterion and contingent reinforcement. It is possible that after exposure to the whole GBG-package, the omission of feedback may produce similar effects on behavior as when feedback is provided.

Wiskow et al. (2019) conducted a more comprehensive study on GBG feedback in a preschool setting. Rather than completing a full component analysis of the GBG, the authors compared four types of feedback; no feedback, visual feedback, vocal feedback or visual + vocal feedback. Using a multi-element design, they demonstrated that either vocal feedback alone or visual + vocal feedback were the most effective in the reduction of disruptive behavior. The GBG with visual feedback alone was also effective in reducing disruptive behavior, but not to the same extent. This was the first study to examine the effectiveness of various feedback during the GBG and provides a basis for further research into different types of feedback during classroom management games, including the GBG and CBGG.

Study Purpose

It is evident that more extensive work is needed on the CBGG in terms of expanding the populations with which it has been used and therefore the first aim of the current study was to examine the CBGG within a mainstream school setting with an Irish adolescent population. The second aim of the current study was to compare the effectiveness of the CBGG with delayed feedback (CBGG-d) and the CBGG with immediate visual feedback (CBGG-i). A version with delayed feedback until the end of class may lead to less distraction to students during the class. This may be preferred by the teacher delivering the intervention as it potentially limits distractions which may be caused by pausing instruction to award points. This said, immediate public posting of feedback throughout the game may offer more reinforcement opportunities for student behavior. Finally, the third aim of the current study was to examine teacher and student acceptability ratings of the CBGG generally, and asked stakeholders to denote a preference for either the CBGG-d or the CBGG-i.

Method

Participants and Setting

Participants were a group of students in a general education secondary school classroom setting, in a densely populated urban area of Dublin, Ireland. The class consisted of 21 participating students (10 females, 10 males, 1 not reported) in their first year of secondary school (approximately equivalent to seventh-grade in the US school system). All participants were of Irish descent and had a mean age of 12.6 (range = 12-14 years). The teacher was a 33-year-old female, general education Mathematics teacher, with five years of teaching experience in the school. She had not used contingencies like the CBGG previously.

Materials

Materials needed for the game included two laminated copies of the class rules, a vibrating timer app for Android/Apple (Tabata Timer: Interval Timer Workout Timer HIIT; Sharafan, 2018) which was installed on the teacher's smartphone, team leader boards (laminated A3 pages) and reinforcers/prizes. Prizes were identified via a preference assessment. These were school cinema passes, school shop tokens, 'positive' journal notes and stationary. The teacher was provided with a checklist which was designed to assist her in keeping on track. Data collectors collected data using paper and pen and intervals were signalled to data collectors through earphones connected to a smart phone.

Dependent Measures and Data Collection

Academically engaged behavior (AEB) and disruptive behavior (DB) were the two dependent measures for which data was collected in this study. Definitions were compiled following consultation with the teacher where she outlined the behaviours most concerning to her, and two preliminary observations of the group during Mathematics instruction.

AEB could be active or passive, therefore definitions for both were devised, however data was collected and collated as one target variable of 'AEB'. Active AEB occurred when the student was actively engaged in the academic task assigned by the teacher, examples of which included reading aloud, copying from the whiteboard or talking to the teacher about

the task. Passive AEB occurred when the student was oriented towards the academic activity, and not engaging in any disruptive behavior, for example looking at the whiteboard while the teacher demonstrates a concept or looking at their copy or text-book while a writing activity was ongoing. A student was not classed as being engaged if their gaze was away from relevant academic material at the time of recording.

Motor and verbal DB were present in the class group. Therefore, both verbal and motor disruption as described here, and are pooled as the target variable, DB. Verbal disruption occurred when a student engaged in any verbalization not authorized by the teacher, for example talking/whispering to a peer, shouting, humming or singing. Motor disruption occurred when a student engaged in movement not related to the assigned academic task for >3 s during the 15 s interval. Examples of motor disruption included being out of seat without permission, turning in their chair, placing their head on the desk or playing with objects in a manner incompatible with the academic task.

Data were collected through a combination of momentary time sampling (AEB) and partial interval recording (DB) four times per week during the 40-min Mathematics class. Sessions lasted 20 min. An individual-fixed method was used; a different student was monitored every 15 s in a predetermined order (Briesch et al., 2015).

Interobserver Agreement (IOA)

Observations were carried out by the first author and trained undergraduate psychology students. IOA data were collected during 25.64% of observation sessions and at least once in each phase. This is in line with the What Works Clearinghouse (WWC) Standards recommendations for collection of IOA during single-case research designs (Kratochwill et al., 2010; WWC, 2017). IOA was calculated using interval-by-interval agreement and dividing the number of agreements by the total number of agreements plus disagreements and multiplying by 100 to obtain a percentage. Mean IOA was 90.56% for DB

(Range= 81.94-100%) and 84.46% for AEB (Range= 72.15-100%). Mean IOA for both of the outcome variables was >80%, which is above the threshold put forward in the WWC Standards. If IOA fell below 80% for any observer, that observer was retrained.

Social Validity

Following the final day of data collection, the teacher and students completed social validity measures. The teacher completed the Intervention Rating Profile (IRP-15; Martens et al., 1985) and students completed a modified version of the Children's Intervention Rating Profile (Mitchell et al., 2015; Witt & Elliott, 1985). The IRP-15 is a rating profile made up of 15 items which assess a teacher's perceptions of intervention acceptability. Items were modified slightly to reflect the application of an intervention to a group of students (e.g., 'child' changed to 'students') and to the present/past tense. Similar changes have been implemented in other research on GBG interventions in group settings (e.g., Mitchell et al., 2015). All items are positively phrased (e.g., "This intervention proved effective in helping to change the problem behavior of the classroom"). Items were rated from 1 (strongly disagree) to 6 (strongly agree), with scores ranging from 15-90. The modified CIRP is a social validity measure with eight items such as 'Did you like participating in the Game?', to which students answered 'yes' or 'no'. Modifications were similar to those implemented in previous research on the GBG with an adolescent population (Mitchell et al., 2015). Mitchell et al. (2015) changed the tense of items from present to past tense, used the term 'students' rather than 'child' and added one item on the rewards used. They also used 'Teamwork Competition' when describing the GBG. In the current study, the intervention was simply referred to as the 'game' in the CIRP. The wording of the final two items were changed from positive wording to negative wording to enhance clarity. Exact items used are outlined later. The highest rating a student could give the game was eight (i.e., eight positive endorsements). Teachers and students were also asked whether they preferred the CBGG-d or the CBGG-i.

Design

A withdrawal design with phases ABACABAC was used to determine the effectiveness of the CBGG-d and CBGG-i. Phase B refers to the CBGG-d and phase C refers to the CBGG-i. Phase changes were determined a-priori. This was essential given the league component of the game; students needed to know the criterion for winning and the day on which they needed to meet that criterion. This was also preferred by the teacher who could plan ahead for when the game was/was not in place.

Procedure

Baseline

Regular classroom instruction took place during baseline and the teacher employed her usual classroom management strategies. No specific reinforcement contingencies were in place for good behavior. DB was addressed by the teacher with verbal warnings, journal notes and sending consistently disruptive students to their form teacher (i.e., head teacher for their class group) or year head (i.e., head teacher for the entire first-year group).

Teacher Training

Teacher training took place during one free class period (35 min). During training, the teacher was provided with an outline for intervention implementation which described both conditions. The researcher showed the teacher how to set up and use the Tabata timer app and they together decided to set the intervals at five min to allow for minimal distraction to the teacher during class time. The teacher and researcher discussed reasonable aims for points needed to obtain the prizes/reinforcers which would be available at the end of each game phase which were adjusted throughout intervention phases based on the students' performance. Time constraints left no time for teacher practice during training, so a feedback phase was implemented which saw the teacher receive immediate feedback from the primary observer across two sessions. This phase is discussed in more detail later.

Intervention: Caught Being Good Game

Following baseline, the teacher introduced the CBGG to the class. The classroom layout allowed for three teams to be assigned based on the seating plan (i.e., there were three columns of students sitting in pairs, so each column was grouped as a team). Particularly disruptive students were dispersed across the classroom, thereby ensuring they were not clustered onto one team. Students were given five min on the first day of the game to choose team names. Class rules were explained, and students were told that their team could earn a point if all members were following these rules when the teacher decided to scan the room.

Upon announcing the game was in play and explaining the rules, the smartphone app timer was started, and the teacher proceeded with her planned tasks. The timer would vibrate every five min, serving as a prompt for the teacher to award points to teams on which all members were following the rules. Each class was 40 min long and the first and last five min were discounted from the game. The timer would vibrate six times in a class. To ensure that the intervals were not exactly fixed and therefore predictable, the teacher was asked to award points at any stage from 0-60 s after the phone vibrated. This was considered as an alternative to using a MotivAider device set on a variable interval schedule (Ford, 2017).

A points criterion for the week was set based on how many days the game would be in place for and on how many points could be earned within that particular time frame. For example, if the game was to be played for six days, the maximum possible points earned for behavior by the end of the final game day would be 36 (6 points x 6 days). The students could also earn an additional 'bonus point' if everyone on their team had their materials for class, leaving the maximum possible points at 42 over 6 days. Therefore, the criterion would be set lower than 42 points. Point criterions were set in conjunction with the class teacher and what could be reasonably attained based on previous behavior. This ranged from 50-75% of the total possible points depending on team progress. For example, if the criterion was set to 30

points for a six-day game-play period and only one of the three teams met this criterion, then the next time the game was played, the criterion would be reduced. Teams meeting or exceeding the criterion by the last day of the game (which was always a Friday) were eligible to gain access to the top prize. This prize was made known to students on the first day of game-play. If a team had not earned enough points by the beginning of Friday's class to earn the top prize, smaller prizes were made available and a smaller, daily criterion put in place.

Feedback Phase (FB). On the first day of game implementation, the teacher implemented the game with 70% treatment integrity. Because of this and given that there had not been an opportunity to practice during training, it was decided to conduct a brief feedback phase. In this phase, the teacher received immediate feedback after class on her implementation integrity. The phase was ceased when treatment integrity reached 80%.

CBGG-d (Phase B). The CBGG-d involved implementation of the game as described. When the timer vibrated, this signalled the teacher to scan the room. Points were recorded discretely in a notebook on her desk. Team totals for the class period were only made known during the final two-three min of class before dismissal. The totals were only then added to a weekly leader-board which was stuck to the whiteboard.

CBGG-i (Phase C). The CBGG-i involved implementation of the game as described. When the teacher was prompted to award points, they were recorded immediately under the team name on the daily leader-board (a laminated A3 page stuck to the whiteboard). The teacher was encouraged to pair the awarding of points in this condition with a positive comment. Points were tallied and added to the weekly leader-board at the end of class.

Treatment Integrity

The teacher checklist included 10 steps for completion of the game (e.g., review the class rules, scan the room for 'rule-following' when signalled). Both the teacher and observer had access to this checklist. The primary observer completed the checklist daily during

intervention sessions. The teacher was asked to keep it on her desk to refer to during game implementation. No specific contingencies were in place if treatment integrity was low (<80%), however the researcher consistently encouraged the teacher in person and via email to use her checklist and ensure each step was completed.

Data Analysis

Data were evaluated visually with the aid of the WWC criteria for evaluation of single-case research designs (Kratochwill et al., 2010; WWC, 2017). Evaluation of design involved looking at number of phases for each condition and number of points per phase. Evidence evaluation involved assessing the changes in level, trend and variability between and across phases. It also involved assessing immediacy of effect, rate of overlap and consistency of data patterns across phases which are the same.

Effect size was calculated using the Tau metric to supplement visual analyses. Tau is an effect size metric based on Kendall's Rank Correlation. In the current study it was calculated based on Tarlow's (2017) recommendations. Tau scores were calculated for each separate AB and AC phase contrast using the Baseline-corrected Tau calculator (Tarlow, 2016). An effect size of .20 may be considered small, .20-.60 moderate, .60-.80 large and .80+ very large (Vannest & Ninci, 2015). Weighted mean effect sizes were calculated for both versions of the game (CBGG-d and CBGG-i) and for both outcome variables (AEB and DB). This was done by weighting effects for each phase transition by their inverse variances (Tarlow, 2017) and calculating a weighted mean effect size using these weights.

Results

Student Behaviors

Visual Analyses

Student AEB and DB across all study phases are illustrated in Figure 1.

Academically Engaged Behavior. During baseline, AEB was low (*M* = 64%, range = 60-80%), and stable after the first data point. During the FB phase, AEB remained low on the first data point while students were introduced to the CBGG for the first time, and increased substantially on the second day the CBGG was in place (*M* = 73%, range = 61-85%). The rates of AEB remained high (*M* = 88.52%, range = 87.5-91.13%) and stable across the continuing CBGG-d phase. The increase here was substantial and there was no overlap with the initial baseline phase.

The first withdrawal phase saw an immediate and large decrease in AEB (*M* = 68.93%, range = 61.25-78.2%) with no overlap between this phase and the preceding intervention phase. The rate of AEB was variable, however pre-treatment levels were reflected in the first and last data points. The CBGG-i was then implemented and AEB increased immediately and substantially (*M* = 82.3%, range = 73.8-86.25%). Behavior remained relatively stable across this phase, however there were two apparent decreases in AEB in the middle of the phase (points 20 and 21). These data points are the only ones to overlap with the preceding withdrawal phase.

The second withdrawal phase saw an immediate and moderate decrease in AEB (*M* = 66.88%, range = 61.25- 72.5%). Reimplementation of the CBGG-d saw AEB increase immediately and remain high and stable across the phase (*M* = 84.69%, range = 81.25-87.5%). This reflects a large change.

During the final withdrawal phase, an immediate and moderate decrease in AEB was evident (*M* = 74.38%, range = 70-82.5%). The CBGG-i was put in place for the final study phase. Immediate increases were not apparent here for AEB (*M* = 82.19% range = 71.25-88.75%). A steady increasing trend for AEB was observed and by the end of the phase, AEB was occurring at a level comparable to the highest AEB in other intervention phases. The overall changes in behavior were smaller than during many of the previous phase changes.

Disruptive Behavior. At baseline, DB was high (*M* = 40.75%, range = 26.25-48.75%). DB did not decrease immediately on the first day of CBGG implementation, but decreased on the second day (*M* = 33.25%, range = 26.25-40.25%). When the FB phase ended and the CBGG-d remained in place for a number of days, DB remained low across the entire phase (*M* = 17.11%, range = 10.13-23.75%). The decrease was substantial when compared to the initial baseline phase, and there was no overlap with the baseline phase.

There was a large increase in DB (*M* = 32%, range = 26.92-35%) when the CBGG-d was withdrawn. There was no overlap with the previous intervention phase. Although DB was variable during this phase, pre-treatment levels were apparent during the first and last data points of the phase. DB decreased immediately and substantially upon introduction of the CBGG-i (*M* = 17.32%, range = 12.5-21.42%). The rate of DB remained stable and low across the phase.

There was an immediate and moderate increase in DB when the CBGG-i was withdrawn (*M* = 29.69%, range = 25-38.75%). The DB trend increased quite steadily during this phase and the rate of DB at the end of the phase was high. When the CBGG-d was reimplemented, DB decreased immediately and relatively steadily across the phase (*M* = 18.13%, range = 12.5-23.75%). Changes in DB were not as large as in AEB in this phase, and some overlap was observed with the preceding withdrawal phase.

The increase in DB was not as pronounced in the final withdrawal phase as between other phases, however the level of DB across the phase remained higher than during intervention phases (*M* = 31%, range = 20-47.5%). Seventy-five percent of the data points for DB did not overlap with the preceding intervention phase. When the CBGG-i was put in place a final time, an immediate decrease was not observed for DB (*M* = 18.44%, range = 13.75-23.75%). A decreasing trend for DB was observed and by the end of the phase, DB was occurring at a similar low rate to other intervention phases.

[Insert Figure 1.]

What Works Clearinghouse Standards

The study design met the WWC standards with reservations rather than fully meeting the standards, as at least three rather than five data points were collected in each phase. The design allowed for at least three attempts to demonstrate an intervention effect for both the CBGG-d and the CBGG-i.

Effect Sizes

Tau coefficients for each phase change are presented in table 1. Effect sizes were large (>.60) across every phase change, except for AEB in the transition from the final withdrawal phase to the CBGG-i, which was a moderate effect size. Weighted mean effect sizes suggest that the CBGG-d was slightly more effective in targeting both AEB and DB than the CBGG-i. Weighted mean effect sizes are large for each intervention and outcome, except for the effect of the CBGG-i on AEB.

[Insert Table 1. Here]

Treatment Integrity

Teacher treatment integrity ranged from 30-100% (*M*=77.5%). There were four steps which were most often missed. The first was when students should have been reminded of how many points their team needed to obtain in order to receive the prize. In addition, the last three steps were commonly missed. These involved announcing the game was finished, announcing and recording teams points and reminding students how many points were now needed in order to get a prize/announce the winners (if it was Friday).

Social Validity

Teacher Rating

The teacher was asked to fill out one questionnaire on the CBGG in general and then to denote a preference for one version over the other. There was also a section on the

questionnaire for written feedback. She scored the CBGG 64 out of 90 on the IRP-15 (*M* = 4.27). She slightly agreed, agreed or strongly agreed with all statements except for two with which she slightly disagreed; "This intervention was reasonable for the problem behavior(s) described" and "This intervention was a good way to handle the behavior problem (s)". The teacher denoted a preference for the CBGG-d, noting that it led to less disruption of class time. She stated that although the intervention was beneficial for most of the students, that some students would need additional behavioral supports and strategies.

Student Rating

Eighteen students completed the modified CIRP following the final day of data collection. Responses to each item are outlined in table 2. The mean score across the respondents was 6.17 (range = 4-8). In general, students enjoyed participation, thought the game was fair and didn't cause them problems, and liked the rewards used. Eleven students preferred the CBGG-i (61.11% of respondents) and six said that they preferred the CBGG-d (33.33% of respondents). One student did not note a preference. Four students provided additional written feedback. Two of these students indicated that they felt sometimes one team member could ruin their chances at a prize (e.g., "I think you should give other people a chance with different groups"). One student gave negative feedback, indicating that they "don't really want to play this game", and one student stated, "I like it".

[Insert Table 2. Here]

Discussion

The primary aim of the current study was to examine the effectiveness of the CBGG in targeting AEB and DB in a first-year class, while comparing delayed with immediate feedback of points during the game. An adolescent student population and their teacher took part and the intervention was examined by monitoring AEB and DB in the group,

implementing a withdrawal design. Behavior improved substantially upon the iteration of the CBGG and behavior returned to or approached baseline levels during each withdrawal phase.

The CBGG-d had large effect sizes for its impact on both AEB and DB. There were immediate and stable increases in AEB and immediate and stable reductions in DB when the CBGG-d was put in place, consistent with previous research by Wahl et al. (2016). Furthermore, the CBGG-d appeared slightly more effective than the CBGG-i in this study which suggests that posting points on the board throughout class may not be an essential component during the game. This finding is consistent with an early study which focused on the GBG rather than the CBGG (Harris & Sherman, 1973). In Harris and Sherman's (1973) component analysis of the GBG, the teacher implemented the game as normal, but withholding feedback by recording it on a page. The current study addressed a key issue with that study by Harris and Sherman (1973) by first implementing the game without feedback before implementation of the game with feedback. This was incorporated into the study design, whereby the first intervention phase was the CBGG-d.

As previously mentioned, although the CBGG-i was effective in targeting AEB and DB, behavior changes were not as pronounced as with the CBGG-d. This was evident in visual analysis and through effect size calculation. The weighted mean effect size for the CBGG-i on AEB was moderate. There are several reasons why this difference may have emerged between versions of the game. First, the immediate delivery of feedback during the CBGG-i may have served to interrupt learning activities and therefore was distracting to the class group. This may have led to lower levels of engagement at specific times during class. This contrasts with feedback delivery during the CBGG-d which did not draw student attention to points during class, thereby not causing distraction. Immediate decreases in DB were apparent during the phases in which the CBGG-i was put in place, however AEB did not increase immediately in the second CBGG-i phase. Increases became apparent across the

first three sessions and stabilised towards the end of the phase. Second, despite the efforts made to provide comparable prizes, it is possible that the prizes available during the CBGG-i may not have been as potent as the prizes used during the CBGG-d. Nonetheless, there was an increasing trend for AEB throughout the CBGG-i phases, with AEB increasing closer to the end of the phase when participants knew that prizes would be awarded. In sum, although the two versions of the CBGG produced increases in AEB and decreases in DB, the changes in behavior were slightly more stable during the CBGG-d. Potential reasons for this stability include less disruption of class time by the teacher for point recording, and more potent prizes. The results also align more generally with previous research which has identified the use of interdependent group contingencies as an evidence-based practice for targeting challenging behavior in the classroom (Maggin et al., 2012).

An ancillary finding emerging from the current research, was that the CBGG was effective using weekly prizes only. Traditionally, the GBG and CBGG have involved the provision of prizes immediately following game-play (e.g., Barrish et al., 1969; Mitchell et al., 2015; Wright & McCurdy, 2012). The class teacher raised concerns over the feasibility of this during a 40 min Mathematics class. The decision was therefore made to trial the game with prizes awarded every four-six days, at the end of a series of classes (a Friday). The CBGG maintained effectiveness across phases, with a weekly goal, rather than a daily goal. Other studies have included both daily and weekly goals (e.g., Lannie & McCurdy, 2007; Wright & McCurdy, 2012), however a weekly goal only had not been examined previously. This serves as a potential avenue for further research in school settings, particularly secondary school settings where students move to different classrooms intermittently throughout the day.

Gauging teacher and student acceptability of the CBGG was a crucially important aspect of this study. By obtaining these ratings and feedback from both stakeholders, issues

can be addressed in future iterations. Teacher and student ratings of the CBGG in general were positive. The class teacher found the intervention acceptable and useful with a score of 64 out of 90, which surpassed Von Brock and Elliott's (1987) suggested target of 52.5 as an indicator of acceptability on the IRP-15. The teacher commented that some students 'needed additional behavior strategies'. This is perhaps reflective of the population sampled where additional supports of many kinds are often needed. Indeed, this was noted by the observers during behavioral observations. Even if the class in general were behaving well, one student could disrupt the rhythm and routine in the classroom. Although systematic direct observation (via partial interval recording and momentary time sampling) provides a reliable approximation of incidences of behavior, not all incidences will have been captured in the behavioral data. For these reasons, the teacher's experiences of the class group that day may not be fully reflected in the quantitative data. Nonetheless if the current social validity data are compared with teacher ratings from a comparable study of secondary school implementation of the GBG (Mitchell et al., 2015), the current ratings compare favorably. The three teachers in the study by Mitchell et al. (2015) rated the GBG 81 (M=5.4), 63 (M=4.2) and 75 (M=5) on the IRP-15 respectively. The current teacher's rating of 64 for the CBGG (M=4.27) was therefore slightly lower than two, and similar to the third secondary school teacher's ratings of the GBG. The teacher denoted a preference for the CBGG-d, stating that it led to less disruption of class time in an already 'volatile classroom environment'. This is not surprising given the disruptive nature of the class. Previous studies have adopted a delayed feedback approach for this reason (Wahl et al., 2016).

Students rated the CBGG positively and most preferred the CBGG-i. Students had likely encountered points provision previously in sports or games, which potentially made the CBGG-i preferable. This finding differs from the teacher's perspective. It is clear that additional research is needed in order to ascertain which version of the game should be put in

place in future with a cohort of this age. The social validity ratings of the CBGG were broadly in line with similar research on both the GBG and the CBGG. This suggests that the CBGG may be a useful intervention across other secondary school classrooms.

Implications for Practice

The current study adds to the literature on games used for classroom management purposes. Specifically, a major implication of the current study is that the CBGG was successful in both formats in the reduction of DB and increase of AEB across a first-year class group. Pending further high-quality investigations, the game may be adopted by secondary school teachers in Ireland as a regular behavior management strategy. Although versions of the GBG had been tested previously with secondary school students, research on the CBGG with this cohort was scarce and a version with delayed feedback had not been tested. The combined alterations to the GBG may be more desirable for use by teachers in secondary school classrooms, as the CBGG is a positive strategy and delayed feedback may lead to less distraction. In secondary school settings, students often move classroom and change subject every 35-40 min. Teachers therefore have a very limited time frame in which to cover course content. The teacher in the current study reported a preference for the CBGG-d over the CBGG-i-, further suggesting that time is of the essence when applying an intervention in a secondary school classroom.

Limitations and Future Directions

A number of limitations must be considered when evaluating the current findings. First, only one class in one school setting were recruited. This limits generalisability of the results to other class groups. It also meant that intervention conditions could not be counterbalanced (i.e., the application of an ACABACAB design with another class group) to buffer against order effects. Future research should determine if it is effective and acceptable with other first-year class groups and perhaps more senior class groups in a secondary school

setting. A second limitation with the design was that only four data points were collected in some phases rather than the five data points recommended in the WWC Standards handbook. The study design therefore met the WWC standards with reservations rather than fully meeting the standards. Importantly, the study meets all other WWC standards. Third, due to the naturalistic setting, some procedural deviations from the intervention protocol arose. It was apparent that the last three steps of the game were most commonly missed by the teacher. This likely relates to time constraints at the end of a class period and the teacher forgetting to implement the last steps. In future, a prompt may be useful towards the end of class to remind the teacher to implement the last few steps of the game. A protocol could also be put in place to counter-act when treatment integrity is low more generally, such as emailed feedback or prompts (e.g., Fallon et al., 2018). Fourth, prizes were provided by the researcher and were purchased with project funds when needed. It may be more beneficial in future to use only prizes which teachers could access for free within the school setting to make the results more applicable to particular school settings. Finally, data were analysed on a group-basis meaning no inferences about individual improvements in behavior could be made.

Conclusion

The current findings suggest that either the immediate or delayed version of the CBGG applied in this study could be considered for use by teachers in lower secondary school settings. This is particularly relevant as schools adopt more positive behavioral approaches, as the CBGG maintains a focus on encouraging rule following rather than punishing rule breaking. It is likely that teachers may prefer using the CBGG-d as it saves some time, however further research is needed on its effectiveness when compared with a version of the game which is more similar to the classic GBG. Future research may focus on replicating the current study findings across two classrooms, counterbalancing conditions and applying the CBGG with older adolescents, such as those preparing for state examinations.

References

Barrish, H. H., Saunders, M., & Wolf, M. M. (1969). Good Behavior Game: Effects of Individual Contingencies for Group Consequences on Disruptive Behavior in a Classroom. *Journal of Applied Behavior Analysis*, *2*(2), 119–124. https://doi.org/10.1901/jaba.1969.2-119

Bowman-Perrott, L., Burke, M. D., Zaini, S., Zhang, N., & Vannest, K. (2016). Promoting Positive Behavior using the Good Behavior Game: A Meta-Analysis of Single-Case Research. *Journal of Positive Behavior Interventions*, *18*(3), 180–190. https://doi.org/10.1177/1098300715592355

Briesch, A. M., Hemphill, E. M., Volpe, R. J., & Daniels, B. (2015). An Evaluation of Observational Methods for Measuring Response to Classwide Intervention. *School Psychology Quarterly*, *30*(1), 37–49. https://doi.org/10.1037/spq0000065

ClassDojo. (2019). https://www.classdojo.com/en-gb/?redirect=true

Conklin, C. G., Kamps, D., & Wills, H. (2017). The Effects of Class-Wide Function-Related Intervention Teams (CW-FIT) on Students' Prosocial Classroom Behaviors. *Journal of Behavioral Education*, *26*(1), 75–100. https://doi.org/10.1007/s10864-016-9252-5

Donaldson, J. M., Wiskow, K. M., & Soto, P. L. (2015). Immediate and distal effects of the good behavior game. *Journal of Applied Behavior Analysis*, *48*(3), 685–689. https://doi.org/10.1002/jaba.229

Fallon, L. M., Collier-Meek, M. A., Kurtz, K. D., & DeFouw, E. R. (2018). Emailed implementation supports to promote treatment integrity: Comparing the effectiveness and acceptability of prompts and performance feedback. *Journal of School Psychology*, *68*, 113–128. https://doi.org/10.1016/j.jsp.2018.03.001

Fishbein, J. E., & Wasik, B. H. (1981). Effect of the Good Behavior Game on Disruptive Library Behavior. *Journal of Applied Behavior Analysis*, *14*(1), 89–93.

https://doi.org/10.1901/jaba.1981.14-89

Flower, A., Mckenna, J. W., Bunuan, R. L., Muething, C. S., & Vega, R. (2014). Effects of the Good Behavior Game on Challenging Behaviors in School Settings. *Review of Educational Research*, *84*(4), 546–571. https://doi.org/10.3102/0034654314536781

Foley, E. A., Dozier, C. L., & Lessor, A. L. (2019). Comparison of Components of the Good Behavior Game in a Preschool Classroom. *Journal of Applied Behavior Analysis*, *52*(1), 84–104. https://doi.org/10.1002/jaba.506

Ford, W. B. (2017). *Evaluation of a Positive Version of the Good Behavior Game Utilizing ClassDojo Technology in Secondary Classrooms* [University of Southern Mississippi]. http://aquila.usm.edu/dissertations

Groves, E. A., & Austin, J. L. (2017). An evaluation of interdependent and independent group contingencies during the good behavior game. *Journal of Applied Behavior Analysis*, *50*(3), 552–566. https://doi.org/10.1002/jaba.393

Harris, V. W., & Sherman, J. A. (1973). Use and analysis of the "Good Behavior Game" to reduce disruptive classroom behavior. *Journal of Applied Behavior Analysis*, *6*(3), 405–417. https://doi.org/10.1901/jaba.1973.6-405

Kleinman, K. E., & Saigh, P. A. (2011). The Effects of the Good Behavior Game on the Conduct of Regular Education New York City High School Students. *Behavior Modification*, *35*(1), 95–105. https://doi.org/10.1177/0145445510392213

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-Case Design Technical Documentation*. https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_scd.pdf

Lannie, A. L., & McCurdy, B. L. (2007). Preventing Disruptive Behavior in the Urban Classroom: Effects of the Good Behavior Game on Student and Teacher Behavior. *Education and Treatment of Children*, *30*(1), 85–98.

https://doi.org/10.1353/etc.2007.0002

Litow, L., & Pumroy, D. K. (1975). A Brief Review of Classroom Group-Oriented Contingencies. *Journal of Applied Behavior Analysis*, *8*(3), 341–347. https://doi.org/10.1901/jaba.1975.8-341

Lynne, S., Radley, K. C., Dart, E. H., Tingstrom, D. H., Barry, C. T., & Lum, J. D. K. (2017). Use of a technology-enhanced version of the good behavior game in an elementary school setting. *Psychology in the Schools*, *54*(9), 1049–1063. https://doi.org/10.1002/pits.22043

Maggin, D. M., Johnson, A. H., Chafouleas, S. M., Ruberto, L. M., & Berggren, M. (2012). A systematic evidence review of school-based group contingency interventions for students with challenging behavior. *Journal of School Psychology*, *50*(5), 625–654. https://doi.org/10.1016/j.jsp.2012.06.001

Martens, B. K., Witt, J. C., Elliott, S. N., & Darveaux, D. X. (1985). Teacher judgments concerning the acceptability of school-based interventions. *Professional Psychology: Research and Practice*, *16*(2), 191–198. https://doi.org/10.1037/0735-7028.16.2.191

Medland, M. B., & Stachnik, T. J. (1972). Good-Behavior Game: A Replication and Systematic Analysis. *Journal of Applied Behavior Analysis*, *5*(1), 45–51. https://doi.org/10.1901/jaba.1972.5-45

Mitchell, R. R., Tingstrom, D. H., Dufrene, B. A., Ford, W. . B., & Sterling, H. E. (2015). The Effects of the Good Behavior Game With General- Education High School Students. *School Psychology Review*, *44*(2), 191–207. https://doi.org/10.17105/spr-14-0063.1

Nolan, J. D., Filter, K. J., & Houlihan, D. (2014). Preliminary report: An application of the Good Behavior Game in the developing nation of Belize. *School Psychology International*, *35*(4), 421–428. https://doi.org/10.1177/0143034313498958

Patrick, C. A., Ward, P., & Crouch, D. W. (1998). Effects of Holding Students Accountable for Social Behaviors during Volleyball Games in Elementary Physical Education. *Journal of Teaching in Physical Education*, *17*, 143–156. https://doi.org/10.1123/jtpe.17.2.143

Pennington, B., & McComas, J. J. (2017). Effects of the good behavior game across classroom contexts. *Journal of Applied Behavior Analysis*, *50*(1), 176–180. https://doi.org/10.1002/jaba.357

Robertshaw, C. S., & Hiebert, H. D. (1973). The Astronaut Game: A Group Contingency Applied to a First Grade Classroom. *School Applications of Learning Theory*, *6*(1), 28–33.

Sharafan, E. (2018). *Tabata Timer: Interval Timer Workout Timer HIIT* (3.6.1). https://play.google.com/store/apps/details?id=com.evgeniysharafan.tabatatimer

Stage, S. A., & Quiroz, D. R. (1997). A Meta-Analysis of Interventions to Decrease Disruptive Classroom Behavior in Public Education Settings. *School Psychology Review*, *26*(3), 333–368.

Tanol, G., Johnson, L., McComas, J., & Cote, E. (2010). Responding to rule violations or rule following: A comparison of two versions of the Good Behavior Game with kindergarten students. *Journal of School Psychology*, *48*(5), 337–355. https://doi.org/10.1016/j.jsp.2010.06.001

Tarlow, K. R. (2016). *Baseline Corrected Tau Calculator*. http://www.ktarlow.com/stats/tau

Tarlow, K. R. (2017). An improved rank correlation effect size statistic for single-case designs: Baseline corrected Tau. *Behavior Modification*, *41*(4), 427–467. https://doi.org/10.1177/0145445516676750

Tingstrom, D. H., Sterling-Turner, H. E., & Wilczynski, S. M. (2006). The Good Behavior Game: 1969-2002. *Behavior Modification*, *30*(2), 225–253.

https://doi.org/10.1177/0145445503261165

Vannest, K. J., & Ninci, J. (2015). Evaluating Intervention Effects in Single-case Research Designs. *Journal of Counseling and Development*, *93*(4), 403–411. https://doi.org/10.1002/jcad.12038

Von Brock, M. B., & Elliott, S. N. (1987). Influence of treatment effectiveness information on the acceptability of classroom interventions. *Journal of School Psychology*, *25*(2), 131–144. https://doi.org/10.1016/0022-4405(87)90022-7

Wahl, E., Hawkins, R. O., Haydon, T., Marsicano, R., & Morrison, J. Q. (2016). Comparing Versions of the Good Behavior Game: Can a Positive Spin Enhance Effectiveness? *Behavior Modification*, *40*(4). https://doi.org/10.1177/0145445516644220

What Works Clearinghouse. (2017). *What Works Clearinghouse™ Standards Handbook (Version 4.0)*. https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf

Wiskow, K. M., Matter, A. L., & Donaldson, J. M. (2019). The Good Behavior Game in Preschool Classrooms: An Evaluation of Feedback. *Journal of Applied Behavior Analysis*, *52*(1), 105–115. https://doi.org/10.1002/jaba.500

Witt, J. C., & Elliott, S. N. (1985). Acceptability of classroom intervention strategies. In T. R. Kratochwill (Ed.), *Advances in School Psychology* (pp. 251–288).

Wright, R. A., & McCurdy, B. L. (2012). Class-Wide Positive Behavior Support and Group Contingencies: Examining a Positive Variation of the Good Behavior Game. *Journal of Positive Behavior Interventions*, *14*(3), 173–180. https://doi.org/10.1177/1098300711421008

Table 1.

Tau Effect Sizes for AEB and DB

	AEB	DB
Baseline to CBGG-d (FB phase not considered)	.775**	-.76**
Withdrawal to CBGG-i	.623*	-.74*
Withdrawal 2 to CBGG-d 2	.756*	-.77*
Withdrawal 3 to CBGG-i 2	.49	-.66
Weighted Mean CBGG-d	.77	-.76
Weighted Mean CBGG-i	.57	-.71

Note. Baseline correction was not required for any of these calculations. *p>.05, **p>.01

Table 2.

Student responses to CIRP items

Statement	% responding 'yes'
Did you like the game used in your classroom?	83.3%
Did you like participating in the game?	77.8%
Do you think other students would like to use the game?	55.6%
Did you like the rewards earned during the game?	88.9%
Do you think the game has helped you do better in Maths class?	50%
Do you think the game was fair?	88.9%
Do you think the game caused any problems for you?	11.1%
Do you think the game caused any problems for your classmates?	16.7%